Computational, biochemical, and NMR-driven structural studies on histone variant H2A.B

Zhang, H.

Chapter 5. The electrostatic potential of the nucleosome acidic patch

This chapter is based on:
Heyi Zhang, Jelmer Eerland, Velten Horn, Hugo van Ingen. The electrostatic potential of the nucleosome acidic patch. *In preparation.*

Contributions of authors:
Amide backbone titration study and histidine side chain titration study were performed by Jelmer Eerland and Velten Horn.

Abstract

The nucleosome surface contains an area with negative electrostatic potential known as the acidic patch, which functions as a binding platform for various proteins to modulate chromatin biology. The dense clustering of acidic residues in this patch may impact their effective pKa and thus the effective electrostatic potential of the acidic patch, which in turn could influence nucleosome-protein interactions. Here, we used solution NMR to determine the structure of the H2A-H2B dimer and its acidic patch free in solution, showing that its core structure, including the acidic patch, agrees well with the structure within the nucleosome. Using pH titration experiments the pKa values for acidic residues as well as histidines that compose or surround the acidic patch were experimentally determined. While for several residues elevated pKa's were predicted based on the structure, the experimental pKa values for glutamate carboxyl group at the acidic patch are all below 4.5. For residue H2A D89, part of the DEE acidic triad, a significantly elevated pKa of 5.6 was observed. Our results establish the order of protonation events in the acidic patch upon lowering pH from physiological values and establish H2B H106 as the first residue to titrate.

Introduction

As the core component of chromatin, nucleosomes serve as a major docking platform for a wide range of proteins that control chromatin biology [1]. Many of these proteins bind to nucleosomes through interaction with a distinct site on the central histone octamer surface, the acidic patch [2]. In addition, nucleosomes can self-interact via this surface, thereby mediating chromatin compaction [3-4]. The acidic patch is formed by the close proximity of six acidic residues from histone H2A and two from H2B, resulting in a defined region of predicted negative electrostatic potential that stands out from the overall positively charged histone surface (**Figure 5.1a, 1b**). Acidic patch binding proteins invariably use one or more arginine side chains to anchor to the acidic patch via hydrogen bonding interactions (**Figure 5.1c**). The proximity of charged residues is known to affect the protonation energies of these titratable residues, possibly resulting in significant shifts in their pKa's [5-6]. In addition, the acidic patch perimeter includes two highly conserved histidine residues (see **Figure 5.1b**) whose protonation degree may be very variable in the physiological range of pH. Thus, a priori it is not clear to what extent residues in and around the acidic patch are protonated under physiological conditions, neither to what extent they might change their protonation state due to subtle changes in local nuclear pH *in vivo*, or due to different buffer pH in experiments performed *in vitro*. Protonation or even fractional protonation of residues in or around the acidic patch may critically alter the surface charge potential and thereby impact the binding affinity of nucleosome-binding proteins. Indeed, unpublished *in vitro* experiments have shown that at least for chromatin factor HMGN2, a change in pH from 7.4 to 6.0 resulted in a ~1000-fold reduced binding affinity to the nucleosome (unpublished data from Hugo van Ingen, Hidenori Kato, Lewis E. Kay, Yawen Bai). Here, we aimed to experimentally determine the pKa values of titratable residues in and around the acidic patch using NMR spectroscopy. NMR spectroscopy is one of the few techniques that can reveal residue-specific side chain pKa's in proteins [5, 8-9]. Changes in the chemical shifts of the nuclear spins in and around the titratable group upon changes in pH are then used to extract pKa values for each

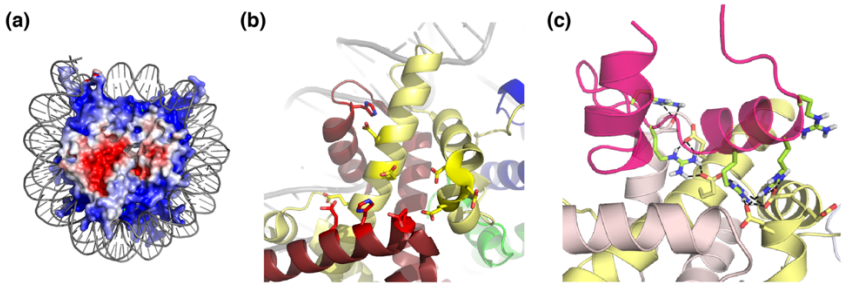

Figure 5.1. Close proximity of acidic residues from histone H2A and H2B forms negatively charged nucleosome surface and functions as arginine anchor of RNF168. (a) APBS predicted electrostatic potential +/- 5 kT on solvent accessible surface of the histone octamer core in nucleosome (PDB id: 2PYO). **(b)** The close proximity of six acidic residues (E55, E60, E63, D89, E90, E91) from H2A and two (E102, E110) from H2B contribute to the acidic patch of the nucleosome. Histidine 46 and 106 from H2B locate around and in the acidic patch are also shown as sticks. H2A is shown in yellow, H2B in red, H3 in blue, H4 in green, and DNA in grey. **(c)** Acidic patch residues function as arginine anchors for the binding of RNF168 on nucleosomes [7].

group. Such methodology has been successfully demonstrated even for very large complexes such as the proteasome using methyl-specific labeling [10], suggesting that this is in principle also possible for the nucleosome. However, this approach relies on the proximity of a methyl-bearing residue to the titratable group of interest. Given the number and the proximity of methyl groups in and around the acidic patch, such indirect readout is unlikely to result in determination of all residue-specific pKa's. We thus resorted to the histone H2A-H2B heterodimer and applied a divide-and-conquer strategy, to be able to map the pKa's of as many titratable groups regardless of the proximity of a methyl-group. We show that in solution the isolated H2A-H2B heterodimer retains the histone-fold core as in the nucleosome and that two glutamate residues in H2A are predicted to have significantly elevated pKa's. By examining pH-dependent chemical shifts changes of both backbone amide and histidine, glutamate, and aspartate side chain groups, we determined the apparent pKa for all titratable groups in and around the acidic patch. Our result show that all acidic residues except H2A D89 have pKa values < 5, H2A D89 has pKa value 5.6, and that pH changes around 6.5 cause large chemical shift perturbations around H2B H106 that lines the acidic patch. These results establish that the glutamate and aspartate residues in the acidic

patch are indeed deprotonated in the physiological range and thus capable to serve as hydrogen bond acceptors for arginine-anchor residues in chromatin binding proteins. Furthermore, the elevated pKa of H2B His106 suggest that it may be fractionally protonated in close-to-physiological conditions and thereby modulate the effective electrostatic potential of the acidic patch and thus the binding affinity of chromatin effector proteins.

Materials and methods

Histone protein production and dimer refolding. *Drosophila melanogaster* (Dm.) canonical histones H2A (Uniprot-id: P84051) and H2B (Uniprot-id: P02283) were expressed in *E. coli* BL21 Rosetta 2 (DE3) cells (Novagen) and purified under denaturing conditions from inclusion bodies by extraction in 6 M guanidine chloride, followed by size-exclusion chromatography in buffer A (7 M urea, 50 mM NaPi pH 7.5, 1 mM EDTA, 150 mM NaCl, 5mM BME) using a Superdex 75 column (GE) and ion exchange with a salt gradient from buffer A to buffer A supplemented with 1 M NaCl. Histones used for NMR studies were produced in D_2O or H_2O-based M9 minimal medium containing desired isotopes. Histone dimers were refolded from equimolar mixes of denatured purified histones by dialysis to 2 M NaCl at room temperature and subsequent purification using size-exclusion chromatography over a Superdex 200 column (GE) in 2M NaCl buffer [11]. Purified dimers were stored at 4 °C before buffer exchange to lower salt concentration for NMR studies.

NMR spectroscopy. The backbone HN, N, Cα, Cβ, C' chemical shifts of H2A and H2B in histone dimer were assigned in previous studies [7, 12] (BMRB accession code 27547 and 27187). The H2A assignment were transferred to low pH conditions (20 mM MES pH5.8, 100 mM NaCl, 10% D_2O) with the aid of triple resonance experiments to confirm the previous assignment and to expand the assignment for the disordered N-terminal tail. Secondary structure of H2A-H2B dimer was predicted by TALOS-N using HN, N, Cα, Cβ, C' chemical shifts [13].

A 3D ^{15}N-edited NOESY with 200 ms mixing time was recorded on a sample containing ~0.5 mM dimer in NMR buffer (20 mM NaPi, 200 mM NaCl, pH 6.5, 5 mM β-mercaptoethanol). The dimer was refolded from perdeuterated ^{15}N-labeled H2B and unlabeled H2A to selectively record intermolecular NOEs.

For pH titration experiments, NMR buffers with pH range from 4.4 to 9.1 were controlled to 200 mM ionic strength with NaCl according to the prescription of the pH calculator website (Rob Beynon, University of Liverpool). 10% D_2O was added to the buffer prior to measurement. Samples were recovered after each titration point and buffer exchanged to next titration point.

For determining backbone amide pKa values, dimers refolded with ^{15}N/^{13}C labeled H2A and unlabeled H2B or ^{15}N labeled H2B and unlabeled H2A were used. HSQC spectra of H2A-H2B dimer (150 μM) in NMR buffers were recorded at 303K on Bruker Avance III HD spectrometers equipped with TCI cryoprobes and operating at 600 or 850 MHz Larmor frequency. Acquisition times was typically 2.5-3 hours per spectrum, with acquisition times of 150 ms for ^{1}H and 100 ms for ^{15}N dimension.

For determining histidine side chain amide pKa values, long-range ^{1}H-^{15}N HMQC experiments[14] were recorded on dimers refolded with ^{15}N-labeled H2B and unlabeled H2A (150 μM) at 303K on Bruker Avance III HD spectrometers equipped with TCI cryoprobes and operating at 600 MHz Larmor frequency.

To enhance sensitivity for detection of glutamate side chains, H2A was made in M9 medium in D_2O, with ^{13}C-protonated glucose, and ^{15}N-labeled NH_4Cl, to remove the alpha protons, and achieve about 50% deuteration on the aliphatic protons. Fractionally deuterated, fully ^{15}N/^{13}C labeled H2A was refolded with unlabeled H2B. The ^{1}H-^{13}C carboxyl projection was recorded for 12h using H(C)CO type experiments to observe side chain carboxyl groups of glutamates and aspartates. Spectra were recorded at 303 K on a Bruker Avance III spectrometer equipped with a cryo-probe and operating at 22.3 T corresponding to 950 MHz ^{1}H Larmor frequency.

To assign side chains of glutamates and aspartates of H2A, several NMR experiments including ^{13}C-constant-time-HSQC, 3D HCCH, 3D CCCONH, and 3D TROSY version of HNCOCACB were recorded at 303 K on a Bruker Avance III spectrometer operating at 21.1 T

corresponding to 900 MHz 1H Larmor frequency, equipped with a cryo-probe. H2A was either fully or fractionally deuterated and fully $^{15}N/^{13}C$ labeled in the dimer sample where H2B was unlabeled.
All NMR data was processed using Bruker Topspin, or NMRPipe and analyzed using NMRFAM-Sparky [15]. Titration data were analyzed using a custom python script written by Hugo van Ingen. Figures of the dimer structure and nucleosome structure (Pdb: 2PYO) were generated in Pymol (The PyMOL Molecular Graphics System, Version 1.7, Schrödinger, LLC).

Structure calculation by CS-Rosetta. Structure calculation of the H2A-H2B dimer was first attempted using CS-Rosetta based only on backbone chemical shifts [16-17]. The HN, N, Cα, Cβ, C' backbone chemical shifts of the core regions of H2A (V26-S97) and H2B (Y34-K122) were used to calculate 3000 structures of the H2A-H2B histone fold core using the CS-Rosetta webserver (https://csrosetta.bmrb.wisc.edu/csrosetta/submit). To allow Rosetta to fold the dimeric core, the two proteins were connected by a random coil (Gly)8 poly-glycine linker into a single chain. The 10 structures with the lowest Rosetta energy model were selected as the best models. However, the final ensemble of CS-Rosetta models was not converged enough to guide the assignment of intermolecular NOEs between H2A and H2B (see below).

Intermolecular NOE assignment by CYANA. ^{15}N-edited NOESY spectrum was recorded on a H2A-[U-^{2}H,^{15}N]-H2B sample. All H2B HN protons were assigned unambiguously based on backbone (HN, N) chemical shifts. Intermolecular NOEs were automatically assigned by CYANA using the atom and residue specific average chemical shift as deposited in the BMRB together with H2A-H2B dimer structure extracted from nucleosome crystal structure (PDB id: 2PYO) as structural reference. About 500 intermolecular NOE cross peaks were observed with chemical shift below 7 ppm. In order to unambiguously assign the large number of intermolecular NOEs observed, cross peaks were grouped and assigned individually according to peak intensities. Briefly, the most intense 35 intermolecular NOE cross peaks together with their potential network support peaks within 0.01 ppm range of each chemical shift were first assigned by CYANA using a tolerance

of 0.4 ppm, which result in 17 unambiguous assignment for the most intense NOEs and their 30 network support NOEs. The assigned atoms were removed from the prot list before the next round of CYANA assignment of median intense NOE peaks together with the unassigned most intense NOE peaks. This second round could not provide unique assignment due to multiple options available for one chemical shift.
The upper distance restraints for the 47 resulted unambiguous intermolecular NOEs were extracted from the .noa output file of CYANA assignment for the most intense peaks. To allow extra freedom of Rosetta structural sampling, an additional 0.5 Å was added to each upper distance restraints before Rosetta calculations.

Structure calculation by Rosetta with NOE distance restraints. The same protocol as described in Chapter 3 was used for calculating and selecting the final structural ensemble for H2A-H2B dimers. Briefly, the distance restraints derived from the intermolecular NOEs were used together with the chemical-shift generated fragments to calculate the structure of the H2A-H2B core region (H2A V26-S97 and H2B Y34-K122 connected by a poly-glycine linker) using the Rosetta AbInitio protocol as described previously (www.rosettacommons.org/demos/latest/public/abinitio_w_chemicalshift_noe/README). In total, 3000 structures were generated and scored on the basis of their the full-atom and distance restraint energy (E) by Rosetta. For each model, backbone chemical shifts of $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$, ^{1}HN and ^{15}N were predicted by SPARTA+. The correspondence between predicted and experimental chemical shifts was used to rescore the initial Rosetta energy to a rescored energy (E'). The model with the lowest E' was selected as the reference model, and the 10 models with the lowest Cα-RMSD from the 20 lowest energy models were selected as the final structure ensemble for the histone fold core. Structural statistics are reported in Supplementary Table S5.1. Predicted pKa's were calculated using PDB2PQR webserver (http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/). Ten structures within the ensemble for H2A-H2B dimer in this study and from previous study were used for pKa predictions to get the average value and error bars. All structure images were created using open source PyMOL.

Results

Isolated H2A-H2B dimer retains the same core structure as in the nucleosome

Previously, the solution structure of the human H2A-H2B dimer was solved by the Nishimura lab based on experimental backbone chemical shifts and the protein folding program ROSETTA, showing considerable variation in the position of H2B αC helix [18]. Since this helix carries two of the acidic residues that contribute to the acidic patch, a potentially dynamic position of the H2B αC helix might affect long-range electrostatic interactions and thus change the electrostatic potential of the acidic patch within the dimer compared to the nucleosome. We thus first wanted to verify that the acidic patch in our experimental system, the fruit fly H2A-H2B heterodimer, retains the same structure as in the nucleosome. Chemical shift indices (CSI) obtained from the experimental Cα and Cβ chemical shifts confirm the presence of the histone fold core secondary structure elements for both histones (**Figure 5.2a**). Similar to the human dimer, the N-terminal α-helix and C-terminal region of H2A are unfolded in solution, while these are defined structural elements in the crystal structure of the nucleosome. Overall, these data confirm Nishimura's observation that the H2A-H2B dimer forms a well-folded histone fold domain with disordered tails in solution.

To obtain the tertiary structure of the histone fold core of the dimer that includes the acidic patch, we used a Rosetta calculation supplemented with sparse distance restraints and backbone chemical shifts . As demonstrated in **Chapter 3** for the H2A.B-H2B variant dimer, addition of distance restraints can be crucial to define the structure of the protein backbone. Indeed, a calculation of the structure purely based on HN, N, Cα, Cβ, C' backbone chemical shifts resulted in an ill-defined histone fold for the 10 models with the lowest Rosetta energy (**Figure S5.1**). Sparse intermolecular distance restraints were extracted from a dedicated NOESY experiment designed to measure intermolecular NOEs between H2B amide protons and H2A backbone or side chain protons, using dimers refolded from unlabeled H2A and perdeuterated and ^{15}N-labeled H2B. For nearly all H2B residues in the core (Y34-K122), including the H2B αC helix, several intermolecular NOEs were observed indicating that the dimers were well-folded (**Figure 5.2b**). Of these, 47 NOEs could be unambiguously assigned using the automated assignment procedure of CYANA [19], taking the nucleosome crystal

structure (PDB id: 2PYO [20]) as the structural reference (see Material and Methods for details) (**Figure 5.2b,c**). Among these NOEs are 7 from the H2B α1 helix, 16 from H2B α2 helix, 14 from H2B αC helix, 7 from H2B loop 1, and 3 from H2B loop2. The dispersion over the core region of H2B suggests that the dimer in solution indeed adopts the same structure as in the nucleosome. Inclusion of these distance restraints in the structure calculation resulted in a well-defined structural ensemble with 1.63 ± 0.29 Å RMSD to the nucleosome dimer structure. (**Figure 5.2d**, see Supplementary **Table S5.1** for Structural statistics).

Acidic patch residues H2A E55 and E60 have high predicted pKa's

Having shown that the H2A-H2B dimer retains the nucleosomal structure in solution, we next submitted the coordinates of the final structural ensemble to the PDB2PQR webserver to predicted the side chain pKa values using the propka protocol. As a control, the predicted pKa from our structure were compared to those predicted from the human dimer structure and those predicted directly from the nucleosomal structure (**Figure 5.3a**). Overall, the three different structures give rise to very similar predicted pKa values. In particular, there is little difference between the values predicted for the fruit fly dimer structure determined here and the less converged human dimer structure. Notably, there are three residues with significantly different predicted pKa in the dimer and nucleosome context. Two of these are in close proximity to the nucleosomal DNA (H2A H81 and H2B D48) which increases their buriedness in the nucleosome and exposes these sidechains to the negative electrical field from the DNA. While the elevated pKa for H2B D48 can be understood in these terms, the reason for the reduced pKa predicted for H2A H81 is unclear. Finally, the predicted pKa for H2A E55 is ~1 unit higher in the nucleosome than in the dimer. Within the nucleosome, the E55 side chain is partly buried in a largely hydrophobic pocket by the H2A the αN helix, the loop to the α1 helix and the H2B α1 helix. The sidechain position is stabilized in this position through hydrogen-bonds to the backbone amide of H2A Q23 and F24 that are part of the loop between the αN and α1 helices. Since the αN helix is not folded in the free dimer, the E55 side chain will be more exposed in the free dimer. We conclude that with exception of H2A E55, H81, and H2B D48, reliable estimates for the

pKa's of residues in and around the acidic patch can be made within the dimer context.

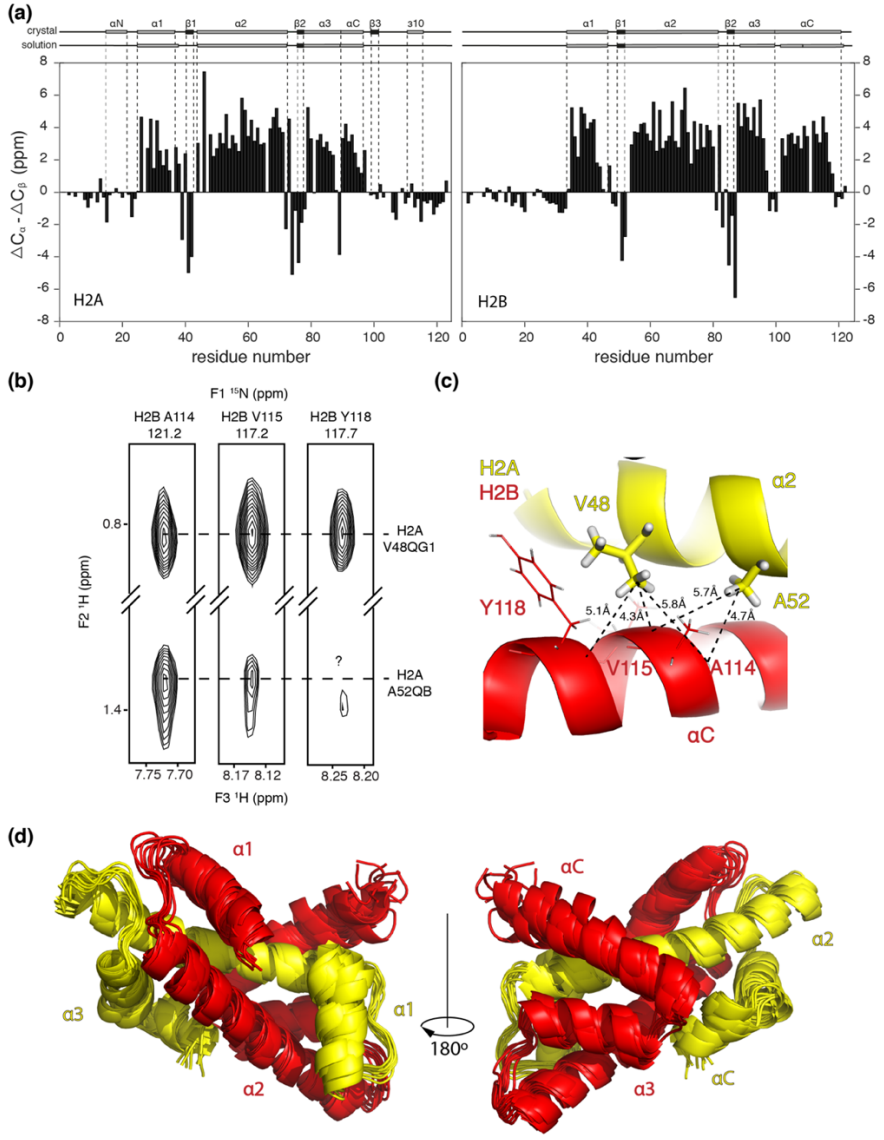

Figure 5.2, isolated H2A-H2B dimer in solution maintains canonical folding as in nucleosomes. (a) secondary structural elements of H2A and H2B in the heterodimer based on experimental chemical shift Cα and Cβ. The secondary structure observed in nucleosome crystal structure (PDB id: 2PYO) and previously reported heterodimer solution structure (PDB id: 2RVQ) are indicated as bars (alpha-helices) and arrows (beta-sheets) above the plot. **(b)** Strips from the 3D ^{15}N-

edited NOESY recorded on heterodimers with perdeuterated ^{15}N-labeled H2B and unlabeled H2A showing intermolecular NOEs between H2B amide and H2A protons. Assignment of H2B amide and H2A resonances are indicated. NOEs shown restrain the H2B αC-helix position. **(c)** Zoom on the H2A-H2B structure (PDB id: 2PYO) corresponding to the intermolecular contact shown in the NOESY spectrum. (d) Superposition of best 10 Rosetta models with lowest rosetta energy and Cα RMSD for H2A-H2B solution structure calculated by Rosetta using backbone chemical shift HN, N, Cα, Cβ, C' and intermolecular NOE distance restraints.

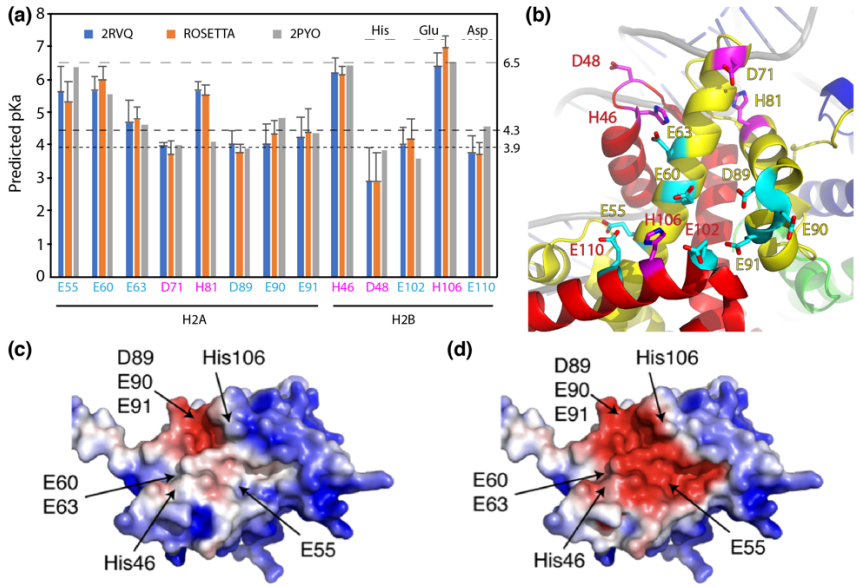

Figure 5.3. Predicted pKa for acidic patch residues represent well with corresponding ones in nucleosome. (a) the predicted pKa values for residues of acidic patch (cyan in (b)) and residues around the nucleosome surface with neutral to negative charges (magenta in (b)) are compared between Nishimura's non-converged solution structure (PDB id: 2RVQ), our converged solution structure calculated by rosetta, and nucleosome crystal structure (PDB id: 2PYO). The 10 models for solution structural ensembles were used for predictions individually and the standard deviations are plotted as error bars. The pKa values for free state Histidine, Aspartic acid and Glutamic acid side chains are indicated as dash lines. **(b)** the residues from plot (a) are highlighted in the nucleosome crystal structure (PDB id: 2PYO), shown as sticks. Nucleosome color coding as H2A is shown in yellow, H2B in red, H3 in blue, H4 in green, and DNA in grey. **(c, d)** Electrostatic surface potential plot of acidic patch in the cases when H2A E55/E60 are neutral and H2B H46/H106 are positively charged (c) and in the case when H2A E55/E60 are negatively charged and H2B H46/H106 are neutral (d)**.** Electrostatic surface potentials were calculated for dimer extracted from PDB id 2PYO using the APBS

plugin in PyMol with the non-linear Poisson-Boltzmann equation contoured at 4 kT/e. Negatively and positively charged surface areas are colored in red and blue, respectively.

Interestingly, three residues have significantly different predicted pKa's compared to their random-coil values. Of these, two residues, H2A E55 and E60, are part of the previously established nucleosome-protein interaction surface in the acidic patch [2, 21](see also **Figure 5.1c**). For both, significantly elevated pKa's are predicted (5.3 for E55 and 6.0 for E60). For E60, this may be due to stabilization of the protonated state by a hydrogen bonding interaction with H2A Y56, whereas for E55 this may due to the partly buried position of this side chain also within the dimer. Within the nucleosome, the pKa of H2A E55 is predicted to even reach 6.4. Together with two histidine's in H2B (H46 and H106) that line the acidic patch, there are thus four residues in and round the acidic patch that may change protonation state depending on buffer pH for the typical pH range used *in vitro*. To illustrate the impact of such changes, we calculated the predicted electrostatic surface potential both for the "worst" case scenario in which all four residues H2A E55, E60 and H2B H46, H106 are all protonated (**Figure 5.3c**), and the "best" case scenario in which these residues are all deprotonated (**Figure 5.3d**). The clear difference in negative electrostatic potential could contribute to loss of binding for proteins that target the acidic patch.

Large changes in the acidic patch chemical environment between pH 6 and 7

We first assessed experimentally the impact of changes in pH on the heterodimer using amide backbone chemical shift perturbation mapping. Using refolded H2A-H2B dimers with uniformly ^{15}N-labeled H2A and unlabeled H2B or vice versa, 2D NH HSQC spectra were recorded under different pH conditions with controlled ionic strength (200 mM). The buffer pH was varied between 4.5 and 9.1 (see **Table S5.2;** using acetate buffers for pH below 5.5 and MES buffers for pH between 5.5 and 6.5, and phosphate buffers for pH above 6.5, and CHES buffer at pH 9). Throughout this pH range well-dispersed spectra were obtained, indicating that the heterodimer remains folded which is in accordance with literature data [22]. Assignments of both H2A and H2B amide resonances were obtained earlier at pH 6.5 [7, 12]

and transferred here to all spectra (see also M&M). Resonances from the disordered histone tails were not visible above pH 8 due to fast exchange of backbone amide proton with the water resonance. Several resonances experience large chemical shift perturbations (CSPs) over the 4.5-9 pH range, including H2A Q23 and E63 as well as H2B H46 and E102. Residues with the largest CSPs cluster in the acidic patch region, also when focusing only on the changes between pH 6 and 8 (**Figure 5.4a**).

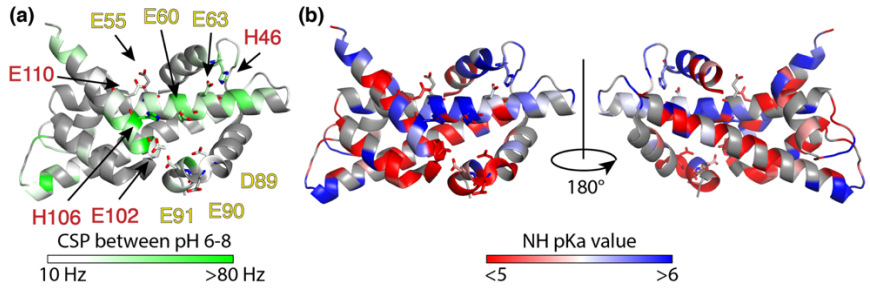

Figure 5.4. Apparent pKa of H2A-H2B dimer core region measured by chemical shift perturbations of backbone amide protons. (a) plot CSP between pH 6 to 8 on crystal structure of H2A-H2B dimer core region (PDB id: 2PYO). Combined CSPs are indicated by color gradient from white (10 Hz) to green (above 80 Hz). Residues that were too overlapped or weak, or with combined CSP less than 10 Hz are colored grey. Acidic patch residues and histidines are indicated by arrows pointing to H2A residues (yellow) and H2B residues (red). **(b)** plot of apparent pKa measured by backbone amide proton on crystal structure of H2A-H2B dimer core region (PDB id: 2PYO). Color coded red to white to blue, represents pKa <5 (red), 5.5 (white), >6 (blue). Residues that were too overlapped or weak are colored grey.H2A acidic patch residues and H2B acidic patch residues together with H46 and H106 are shown as sticks.

As the amide backbone chemical shifts are not only sensitive to electrostatics, but even more so to changes in hydrogen bonding strength and local conformation, the observed CSPs can only be taken as indirect reporters of changes in protonation state. Moreover, sensitivity to multiple (nearby) protonation events can result in curved peak trajectories, that can complicate the analysis. Fitting of the CSP-derived pH titration profiles to a modified version of the Henderson-Hasselbalch equation thus results in a single, residue-specific apparent pKa, even for residues without titratable group and even in the case where multiple protonation events are sensed (see **Table S5.3)**. For

example, H2A residues A65 displays a clear and linear pH –dependent CSP that can be fitted to an apparent pKa value of 5.75 which may be due to protonation of the nearby H2B H46 side chain at low pH (**Figure 5.5a**). An extremely curved trajectory was observed for H2A E63 that is near to both H2A E63 and H2B H46 side chains, and thus may be sensitive to two protonation events (**Figure 5.5d**).

In cases where the amide proton is hydrogen bonded to a titratable group, the apparent pKa can be taken to reflect directly pKa of the side chain involved. For example, one of the largest CSPs is observed for H2B E102 amide which can be explained by a change in hydrogen bonding to the nearby carboxyl group of H2A E91 upon lowering the pH below ~5.3 (**Figure 5.5b**). The resulting best-fit apparent pKa of 4.28 thus likely represents the pKa of the H2A E91 carboxyl group. Similarly, the fitted value obtained for the H2B L103 amide likely reflects the E90 pKa.

Interestingly, the pKa of H2A E55 can be estimated to be 4.5, based on the apparent pKa's for the H2A Q23 and F24 amides (**Figure 5.5c**). As noted above, these amides are hydrogen bonded to the E55 carboxyl group in the nucleosome. While the preceding αN helix is not formed within the free dimer and the loop containing these amides is likely more flexible in the dimer, the strongly downfield shifted $^{1}H_{N}$ chemical shifts suggest these remain hydrogen bonded in the dimer. Furthermore, H2A E55 is the only titratable group in close proximity to these amides. This estimate for the E55 pKa value is considerably lower than the structure-based predicted value of 5.3 and more in line with the expected values for glutamate side chains.

The last acidic patch residue for which the backbone data can be used to derive the sidechain pKa is H2A D89. Its sidechain is hydrogen bonded to the E91 and L92 backbone amides. Surprisingly, these report an apparent pKa of 5.6 (**Fig 5e**, significantly higher than the random coil (3.8) or structure-predicted pKa values (3.9).

When plotted on the structure, inspection of the residue-specific pKa shows that most residues in and around the acidic patch, in particular near the H2A D89, E90, E91 triad, report apparent pKa values below 5 (**Figure 5.4b**). The opposite end of the acidic patch, near H2A E55 and H2B E102, as well as the center around H2A E60 and E63 have apparent pKa higher than 5. These are relatively close to histidine H106 and H46 from H2B, that line the acidic patch. These observations,

coupled with the large CSPs observed for the acidic patch, suggest that the effective electrostatics of the acidic patch may indeed change in the pH range of 6-8. This change is likely due to changes in protonation state of the H2B histidines H106 and H46, and H2A D89.

Direct pKa measurement of Glu and Asp by side chain carboxyl carbon chemical shifts

To verify whether the acidic patch glutamates or aspartates have indeed low pKa as suggested by the analysis above, we next wanted to observe the chemical shifts of the carboxyl carbons directly as they are more unambiguous and sensitive reporters of the protonation state. These ^{13}C chemical shifts typically move ~4 ppm upfield as the carboxyl group undergoes protonation. Using dimer (25 kDa) with fractionally deuterated and ^{13}C labeled H2A, 2D CH correlation spectra were recorded using a HCCO pulse sequence where each correlation connects the $^{13}C\gamma/C\delta$ carboxyl groups to the $H\beta/H\gamma$ of aspartate or glutamate [23]. Here only H2A was isotope labeled, since H2A contains most of the relevant acidic residues, including E60 which may have elevated pKa according to the structure based pKa prediction and backbone based pH titration. Histone H2A contains 7 glutamate and 2 aspartate acids together with 7 asparagines and 4 glutamines, which in principle should generate 40 observable signals in these spectra. Signals from Asn/Asp and Gln/Glu can be discriminated based on distinct chemical shifts for the carboxyl carbon (176.4/177.8 ppm for $C\gamma$ of Asn/Asp, 179.3/181.2 ppm for $C\delta$ of Gln/Glu) and for proton $H\beta/H\gamma$ (~2.7 ppm for Asn/Asp, ~2.2 ppm for Gln/Glu). The 2D HCO spectrum shows several peaks in the glutamate region, which are unfortunately heavily overlapped (**Figure 5.6c**). The Gln and Asn residues are also visible from the 2D HCO spectra and well separated from Glu residues. No CSP were observed during the titration for Gln and Asn residues (**Figure 5.6b**). Unfortunately, signals for the 2 Asp residues could not be detected, likely due to lack of more extensive line broadening for these shorter sidechains [23]. To assign the glutamate $C\delta/H\gamma$ resonance, we attempted to correlate the $H\gamma$ resonances to the assigned backbone resonances via both triple resonance backbone and sidechain resonances. However, neither strategy was successful due to poor sensitivity. The corresponding region in the CH constant time HSQC spectrum shows 7 resonance pairs that likely correspond to the

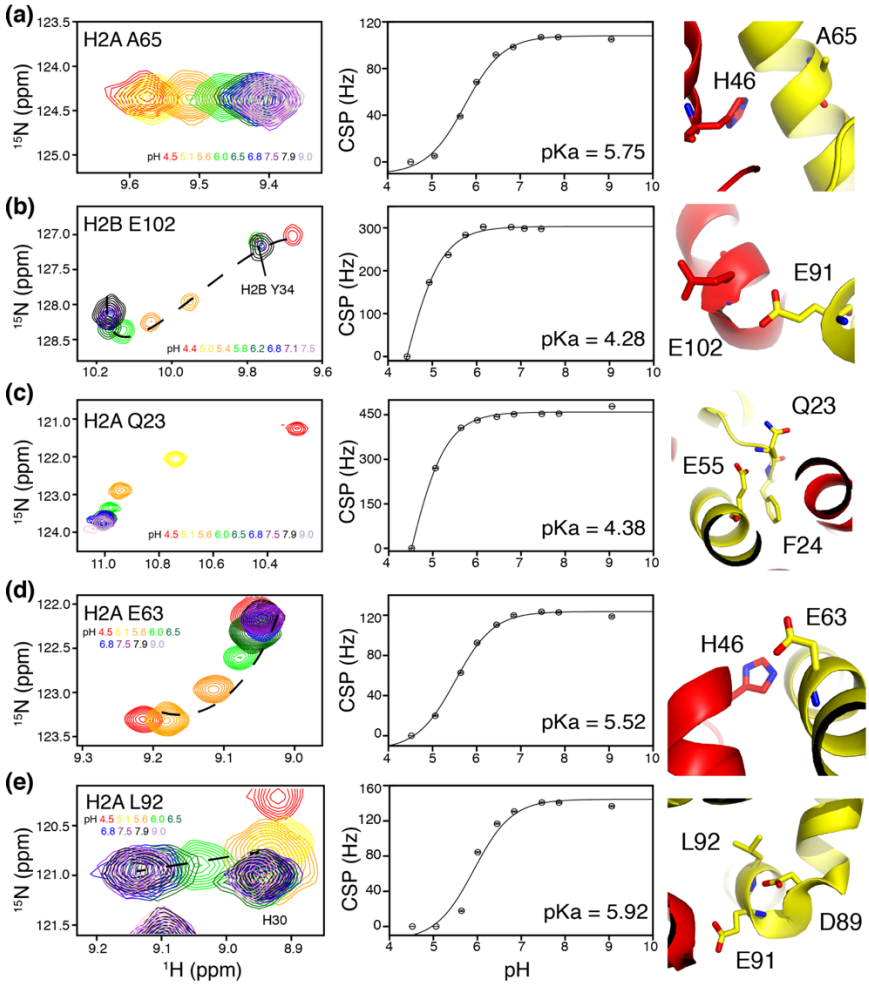

Figure 5.5. Apparent pKa of H2A-H2B dimer core region measured by chemical shift perturbations of backbone amide protons. (a-e) CSP of residues are shown on the left column, titration curves fit combined CSP to pH and the apparent pKa determined are shown in the middle column, and the structural references using the crystal structure of H2A-H2B dimer (PDB id: 2PYO) are shown on the right column.

seven Glu residues in H2A (**Figure 5.6a**). Due to the high dynamics of the C-terminal H2A tail, we could tentatively assign C-terminal E120 to the most intense pair of signals. By comparing the CH constant time HSQC spectrum with the 2D HCO spectrum at similar pH conditions (5.8 and 6.0 respectively), we could arbitrarily assign the other six

glutamates to the overlapped peaks (**Figure 5.6c**). By following these arbitrarily assigned glutamate peaks in the 2D HCO spectrum at titrated pH conditions, valuable information of the pKa values of the observed resonances was extracted (**Figure 5.6d**). During titration, all peaks shift more or less together, indicating the corresponding glutamate side chains have similar pKa values. All glutamates side chain carboxyl group pKa are below 5 (**Figure 5.6e**). Since the dimer does not fold stably in conditions with pH lower than 4.5, no data could be collected to reach the fully protonated state. Nevertheless, the limited CSPs (~1.1 ppm) and approximate fit to the Henderson-Hasselbalch equation clearly indicate their pKa value is lower than 5. These findings thus indicate that the acidic residues in the acidic patch have pKa's close to their default values and thus point to the histidines as being responsible for the large CSPs in the acidic patch between pH 6 and 8.

pKa values of H2B H106 and H46

To experimentally determine the pKa values of the histidines that line the acidic patch, we used heterodimers with $^{15}N,^{13}C$-labeled H2B and unlabeled H2A and recorded HMQC spectra probing long-range correlation between the sidechain $^{1}H\varepsilon1$, $^{1}H\delta2$, $^{15}N\varepsilon2$ and $^{15}N\delta1$ spins in the histidine side chains. Since H2B contains both H46 and H106 that line the acidic patch, as well as an additional histidine H79 close to the dimer-dimer interface, a titration experiment on this single sample suffices to obtained the relevant pKa's. The spectra indeed show correlations for each of the three histidines in H2B (**Figure S5.2**). Titration of the pH between 4.4 and 7.5, resulted in clear chemical shift changes (**Figure 5.7a**) that could be fitted to pKa's of 5.9, 6.5 and 6.7 (**Figure 5.7b**). Although side chain assignment of the histidines is currently lacking, we could assign the three histidines based on comparison of the side chain determined pKa data with the backbone amide proton determined pKa's. Histidine number 3 has the largest chemical shift perturbation in the proton dimension and its pKa is determined to be below 6 (**Figure 5.7a,b**). In the dimer structure, H2B H46 side chain is close to several H2A residue backbone amide protons whose pKa values were determined to be below 6 (H2A E63-N67, see **Table S5.2**), while residues around H2B H79 have backbone amide proton pKa value above 6.5 (H2A A39 6.7, H2B R76 6.8, Y80 6.5).

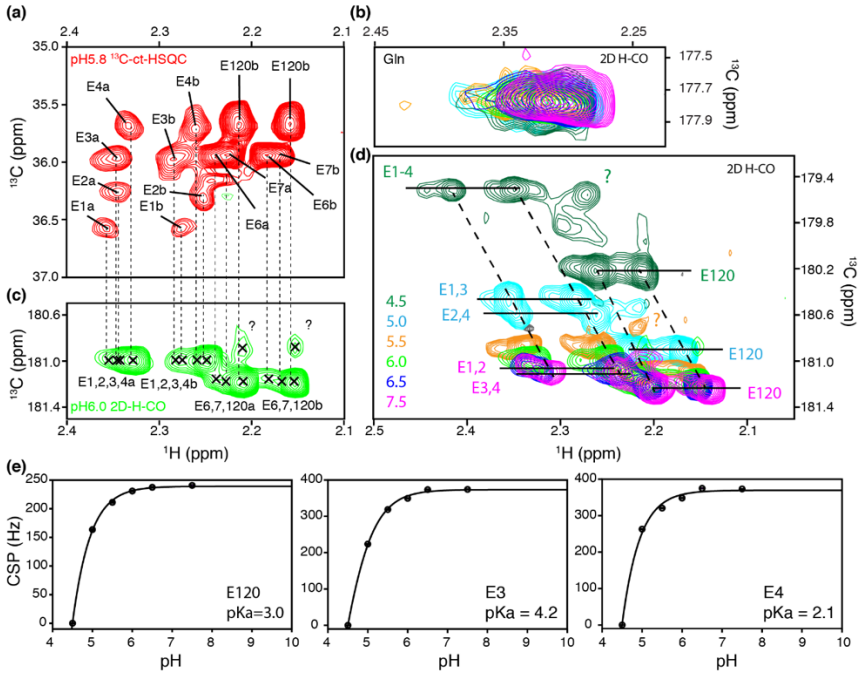

Figure 5.6. glutamates in H2A has pka below 5 for side chain carboxyl group in dimer. (a) ^{13}C-H constant time HSQC spectrum of H2A glutamates show well resolved peaks for 7 glutamates in sequence. Each glutamate has a duplicated peak for the side chain carboxyl group. **(b, c, d)** H-CO 2D spectra measured using HCACO sequence centered on side chain. Glutamines (b) has different chemical shifts than glutamates (c, d) on 2D H-CO spectra. (c) Glutamates side chain carboxyl group heavily overlap with each other. Dash lines showing the assignment of 7 glutamates based on ^{13}C-H constant time HSQC spectrum. Question marked peaks are not traceable at other pH conditions. (d) CSPs of H2A glutamates shown by 2D H-CO spectra overlay at different pH conditions. During pH titration, obvious shift started to occur when pH drops to 5. Peak positions are marked at representative pH conditions. Arbitrarily assigned E6 and E7 shifts closely represent the pattern of E120. **(e)** titration curves fit CSP to pH are shown for H2A E120, arbitrarily assigned E3 and E4.

For histidine number 2, we have observed a change in shifting direction of 2Nε2-Hδ2 and 2Nδ1-Hδ2 side chain group when the pH drops below 5 (**Figure 5.7e**). Similarly, changes in shifting direction of backbone amide proton for H2B A107 and L103 were observed (**Figure 5.7f,g**). Since H2B A107 and L103 is close to H106 in the dimer structure, we can confirm the assignment for histidine number 2 to H106. Thus, we assign the measured pKa for number 1 to H79,

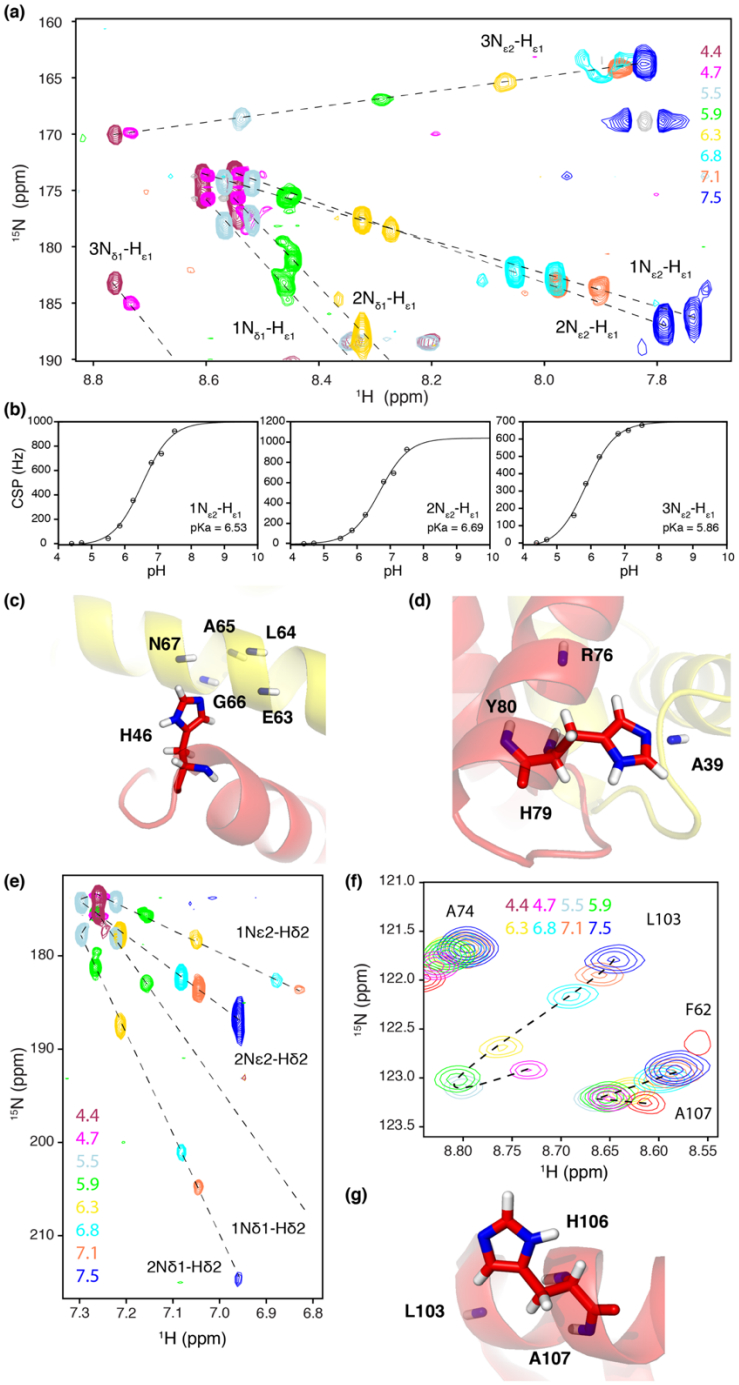
(a)
3Nε2-Hε1
4.4
4.7
5.5
5.9
6.3
6.8
7.1
7.5
3Nδ1-Hε1
1Nε2-Hε1
2Nδ1-Hε1
1Nδ1-Hε1
2Nε2-Hε1
15N (ppm)
1H (ppm)
(b)
CSP (Hz)
1Nε2-Hε1
pKa = 6.53
2Nε2-Hε1
pKa = 6.69
3Nε2-Hε1
pKa = 5.86
pH
(c)
N67
A65
L64
G66
E63
H46
(d)
R76
Y80
A39
H79
(e)
1Nε2-Hδ2
2Nε2-Hδ2
1Nδ1-Hδ2
2Nδ1-Hδ2
4.4
4.7
5.5
5.9
6.3
6.8
7.1
7.5
15N (ppm)
1H (ppm)
(f)
A74
4.4 4.7 5.5 5.9
6.3 6.8 7.1 7.5
L103
F62
A107
15N (ppm)
1H (ppm)
(g)
H106
L103
A107

Figure 5.7. pka value determined by histidine side chain. (a) CSP of 3 histidines side chain amide groups from H2B at different pH conditions. Peak shifting directions are marked as dashed lines. **(b)** titration curves fit CSP to pH are shown for each histidine. **(c)** Backbone amide protons from H2A E63-N67 close to H2B H46 side chain amide group report apparent pKa's below 6 (see table S5.2). **(d)** backbone amide groups from H2B R76 and Y80 and H2A A39 close to H2B H79 side chain amide group report apparent pKa's above 6.5 (see table S5.2). **(e)** change of CSP direction at pH below 5 of histidine's side chain amide groups of 2Nε2-Hδ2 and 2Nδ1-Hδ2. **(f)** change of CSP direction at pH below 5 of backbone amid groups of H2B L103 and A107. **(g)** H2B L103, H106 and A107 are close to each other in dimer structure (PDB id: 2PYO).

histidine number 2 to H2B H106, and histidine number 3 to H2B H46. To conclude, we have measured histidine side chain pKa values of H2B histidines as 5.9 for H46, 6.5 for H79 and 6.7 for H106.

Discussion

The nucleosome acidic patch is a binding surface for many nucleosome remodeling factors. The electrostatic potential of the surface is essential for binding of specific partners, and moderate changes may cause binding deficiency. For chromatin factor HMGN2 it was found that a drop in buffer pH from 7 to 6 results in a drastic decrease in binding affinity to nucleosome (unpublished data from Hugo van Ingen, Hidenori Kato, Lewis E. Kay, Yawen Bai). Such peculiar pH dependent binding affinity may directly be caused by changes in the electrostatic potential of nucleosome surface. Here, we sought to experimentally determine the pKa values of all titratable groups in and around the acidic patch within the free H2A-H2B dimer.

We first determined the H2A-H2B solution structure using a CS-ROSETTA protocol supplemented with sparse intermolecular distance restraints. The resulting structure is well defined and close to nucleosomal dimer structure. Notably, this structure does not show evidence for an intrinsic conformational heterogeneity present in the dimer as was argued in a previous study using CS-ROSETTA only[18].

Determination of all side chain pKa was complicated by the substantial size of the H2A-H2B dimer (25 kDa). While residue specific mapping of sidechain pKa has been demonstrated before for smaller protein domains and intrinsically disordered proteins, the enhanced relaxation in this system made it challenging to directly probe the protonation of

all sidechain carboxyl groups in an unambiguous manner. Yet, combination of these data with apparent pKa derived from backbone amide data and experimental pKa's for the histidine sidechains allows to assign pKa values or estimated values to all titratable groups in and around the acidic patch (**Table 5.1**). Our results show that all glutamates residues in the acidic patch have pKa below 5, and thus unambiguously falsify the high predicted pKa for H2A E55 and E60. Surprisingly, for the sole aspartate in the acidic patch, H2A D89, we find a best estimate of 5.6 for its pKa. This estimate is based on the backbone data and could unfortunately not be verified directly in sidechain experiments due to too low a sensitivity. For two histidines that line the acidic patch, H106 and H46, pKa values of 6.7 (H106) and 5.8 (H46) were determined, respectively. At physiological pH, the acidic patch is thus highly negatively charged. We found that reduction of the pH below the physiological value results in large CSPs in and around the acidic patch. Our data indicate that this is first due to the protonation of H2B H106, then due to protonation of H2B H46 and H2A D98. Protonation of these residues likely alters their sidechain position, which in particular for the aromatic histidine rings will result in large chemical shift changes for surrounding backbone amides.
In situ pH measurements have indicated that the nuclear pH is 7.8, slightly higher than the cytosolic pH (7.4) [24]. At pH of 7.8, the histidine H106 and H46 are for 7.4% and 1.0% in their protonated, positively charged form, while D89 is <0.6% in the neutral state. At pH 7.4 this increases to 16.6%, 2.5% and 1.5%, respectively. The reduced pKa for H46 may help to retain a pronounced acidic character near zone I of the acidic patch (H2A E60, L64, D89, E91, and H2B E102, L103) which is the key interaction site of nucleosome-binding proteins [2, 21]. Such slight changes in protonation levels are unlikely to result in significant perturbation of binding of chromatin factors to the acidic patch. In addition, burial of these residues upon protein binding may reduce their pKa's by stabilizing the deprotonated state. Yet, it is conceivable that changes in the local microenvironment could result in more pronounced acidification that would disfavor binding.

Table 5.1. Predicted pKa and experimentally measured pKa values for all titratable groups in H2A-H2B dimer.

residue	Predicted pKa (dimer / nucleosome)	Best estimated from backbone N-H apparent pKa	Best estimated from side chain pKa	Final best estimated pKa
H2A H30	5.68/1.28	4.36		
H2A E40	3.82/4.84	5.91	<5	<5
H2A E55 (AP)	5.30/6.37	4.53 (hb to Q23, F24)	<5	4.53
H2A E60 (AP)	5.99/5.56	6.15	<5	<5
H2A E63 (AP)	4.78/4.62	5.52	<5	<5
H2A D71	3.73/4.02	4.58		
H2A H81	5.53/4.10	N/A (hb to I78)		
H2A D89 (AP)	3.79/3.88	5.59 (hb to E91, L92)		5.59
H2A E90 (AP)	4.33/4.82	4.24	<5	<5
H2A E91 (AP)	4.40/4.37	4.35 (hb to H2B E102, L103)	<5	<5
H2A E120	-/5.08	4.18	<5	4.18
H2A A123 (C-termini)	-/3.24	4.08		4.08
H2B E32	-/3.87	6.10		
H2B H46	6.14/6.43	5.68 (close to H2A G66, N67)	5.86	5.77
H2B D48	2.88/3.85	5.83		
H2B D65	4.42/2.85	-		
H2B E68	4.50/6.28	-		

H2B E73	4.25/3.78	-		
H2B H79	6.55/5.23	6.91	6.53	6.53
H2B E90	3.32/4.02	3.77 (hb to T87)		3.77
H2B E102 (AP)	4.20/3.59	4.50		
H2B H106	6.96/6.53	6.61	6.69	6.69
H2B E110 (AP)	3.72/4.55	5.14		
H2B K122 (C-terminus)	2.43/3.07	4.25		4.25

Comparison of the experimental and predicted pKa values highlights the challenge in accurate prediction of pKa values and the need for experimental validation (**Table 5.1**). The predictions indicate that for residues close to the nucleosomal DNA, e.g. D48 or H81, the pKa measured may be altered in the nucleosome. This may also be the case for H2A E55, for which a higher pKa values was predicted in the nucleosome context. Assuming that the experimental value in the nucleosome is higher by one unit, as in the predictions, this would result in a pKa of ~5.5 for E55 in the nucleosome. This residue would thus be the fourth residue to titrate upon dropping the pH below physiological value, after H106, H46 and D98.

In conclusion, we have determined the solution structure of the H2A-H2B dimer and studied the pH titration of the acidic patch residues. We find that the acidic patch is indeed highly acidic at physiological pH and remains so down to pH ~6.7 at which point H2B H106 starts to titrate. Further reduction in pH will cause protonation of both H2B H46 and H2A D98 in that will significantly decrease the electronegativity near the key protein interaction site of the acidic patch. In addition, H2A E55 may also get significantly protonated around pH 5.5 when in the nucleosome context. These results are important for proper planning of *in vitro* experiments as well as for accurate molecular dynamics simulations of nucleosome-nucleosome or nucleosome-proteins interactions that are mediated through the acidic patch.

References

1. Speranzini, V.; Pilotto, S.; Sixma, T. K.; Mattevi, A., Touch, act and go: landing and operating on nucleosomes. *EMBO J* **2016,** *35* (4), 376-88.

2. McGinty, R. K.; Tan, S., Recognition of the nucleosome by chromatin factors and enzymes. *Curr Opin Struct Biol* **2016,** *37*, 54-61.

3. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **1997,** *389* (6648), 251-60.

4. Chodaparambil, J. V.; Barbera, A. J.; Lu, X.; Kaye, K. M.; Hansen, J. C.; Luger, K., A charged and contoured surface on the nucleosome regulates chromatin compaction. *Nat Struct Mol Biol* **2007,** *14* (11), 1105-7.

5. Pahari, S.; Sun, L.; Alexov, E., PKAD: a database of experimentally measured pKa values of ionizable groups in proteins. *Database (Oxford)* **2019,** *2019*.

6. Bombarda, E.; Ullmann, G. M., pH-dependent pKa values in proteins--a theoretical analysis of protonation energies with practical consequences for enzymatic reactions. *J Phys Chem B* **2010,** *114* (5), 1994-2003.

7. Horn, V.; Uckelmann, M.; Zhang, H.; Eerland, J.; Aarsman, I.; le Paige, U. B.; Davidovich, C.; Sixma, T. K.; van Ingen, H., Structural basis of specific H2A K13/K15 ubiquitination by RNF168. *Nat Commun* **2019,** *10* (1), 1751.

8. Oda, Y.; Yamazaki, T.; Nagayama, K.; Kanaya, S.; Kuroda, Y.; Nakamura, H., Individual ionization constants of all the carboxyl groups in ribonuclease HI from Escherichia coli determined by NMR. *Biochemistry* **1994,** *33* (17), 5275-84.

9. Platzer, G.; Okon, M.; McIntosh, L. P., pH-dependent random coil (1)H, (13)C, and (15)N chemical shifts of the ionizable amino acids: a guide for protein pK a measurements. *J Biomol NMR* **2014,** *60* (2-3), 109-29.

10. Velyvis, A.; Kay, L. E., Measurement of active site ionization equilibria in the 670 kDa proteasome core particle using methyl-TROSY NMR. *J Am Chem Soc* **2013,** *135* (25), 9259-62.

11. Luger, K.; Rechsteiner, T. J.; Richmond, T. J., Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* **1999,** *304*, 3-19.

12. Corbeski, I.; Dolinar, K.; Wienk, H.; Boelens, R.; van Ingen, H., DNA repair factor APLF acts as a H2A-H2B histone chaperone through binding its DNA interaction surface. *Nucleic Acids Res* **2018,** *46* (14), 7138-7152.

13. Shen, Y.; Bax, A., Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular NMR* **2013,** *56* (3), 227-241.

14. Pelton, J. G.; Torchia, D. A.; Meadow, N. D.; Roseman, S., Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from Escherichia coli, using two-dimensional heteronuclear NMR techniques. *Protein Sci* **1993,** *2* (4), 543-58.

15. Lee, W.; Tonelli, M.; Markley, J. L., NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015,** *31* (8), 1325-7.

16. Shen, Y.; Bryan, P. N.; He, Y.; Orban, J.; Baker, D.; Bax, A., De novo structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci* **2010,** *19* (2), 349-56.

17. Shen, Y.; Vernon, R.; Baker, D.; Bax, A., De novo protein structure generation from incomplete chemical shift assignments. *J Biomol NMR* **2009,** *43* (2), 63-78.

18. Moriwaki, Y.; Yamane, T.; Ohtomo, H.; Ikeguchi, M.; Kurita, J.; Sato, M.; Nagadoi, A.; Shimojo, H.; Nishimura, Y., Solution structure of the isolated histone H2A-H2B heterodimer. *Sci Rep* **2016,** *6*, 24999.

19. Guntert, P.; Buchner, L., Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* **2015,** *62* (4), 453-71.

20. Clapier, C. R.; Chakravarthy, S.; Petosa, C.; Fernandez-Tornero, C.; Luger, K.; Muller, C. W., Structure of the Drosophila nucleosome core particle highlights evolutionary constraints on the H2A-H2B histone dimer. *Proteins* **2008,** *71* (1), 1-7.

21. Fang, Q.; Chen, P.; Wang, M.; Fang, J.; Yang, N.; Li, G.; Xu, R. M., Human cytomegalovirus IE1 protein alters the higher-order chromatin structure by targeting the acidic patch of the nucleosome. *Elife* **2016,** *5*.

22. Karantza, V.; Baxevanis, A. D.; Freire, E.; Moudrianakis, E. N., Thermodynamic studies of the core histones: ionic strength and pH dependence of H2A-H2B dimer stability. *Biochemistry* **1995,** *34* (17), 5988-96.

23. Everill, P.; Sudmeier, J. L.; Bachovchin, W. W., Direct NMR observation and pKa determination of the Asp102 side chain in a serine protease. *J Am Chem Soc* **2012,** *134* (4), 2348-54.

24. Seksek, O.; Bolard, J., Nuclear pH gradient in mammalian cells revealed by laser microspectrofluorimetry. *J Cell Sci* **1996,** *109 (Pt 1)*, 257-62.

Supplements

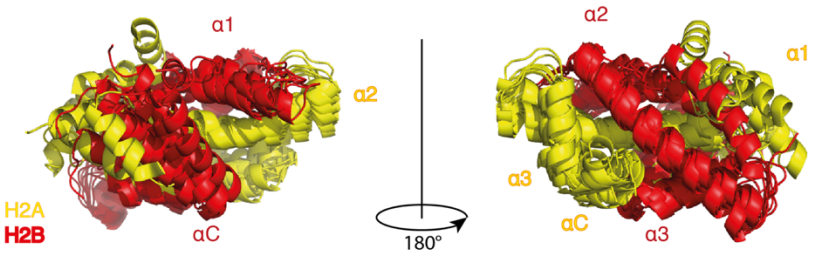

Figure S5.1. Superposition of the 10 CS-Rosetta models with the lowest Rosetta energy to align the H2A and H2B α2 helices, showing ill-defined positions of H2A α1 and H2B αC helices within the dimer structure.

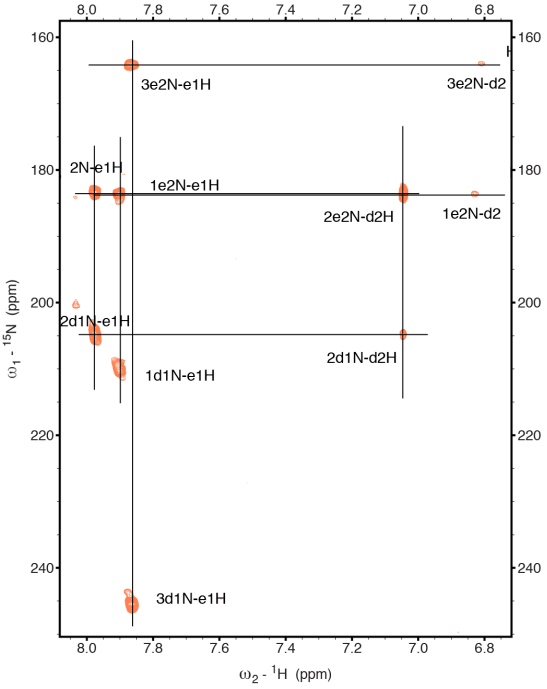

Figure S5.2. Side chain H-N spectra for three histidine's in H2B at pH 7.1.

Table S5.1. Structural statistics for the core region (H2A V26-S97 and H2B Y34-K122) of isolated H2A-H2B heterodimer

A. Restraint information	
number of intermolecular NOEs	47
B. Average RMS deviation from experimental restraints	
All experimental distance restraints (Å)	0.3 ± 0.5
C. Coordinate RMS deviation[a] (Å)	
Average RMSD to mean	
Ordered backbone atoms	1.1 ± 0.2
Ordered heavy atoms	1.5 ± 0.2
Global backbone atoms	1.2 ± 0.2
Global all heavy atoms	1.6 ± 0.2
Average RMSD to 2PYO	
Ordered backbone atoms	1.6 ± 0.3
Ordered heavy atoms	2.3 ± 0.3
Global backbone atoms	1.9 ± 0.3
Global all heavy atoms	2.5 + 0.3
Average Pairwise RMSD	
Ordered backbone atoms	1.5 ± 0.3
Ordered all heavy atoms	2.2 ± 0.3
Global backbone atoms	1.8 ± 0.3
Global all heavy atoms	2.3 ± 0.3

[a] Ordered regions correspond to all helices at the core

Table S5.2: NMR samples and buffers (with ionic strength of 200mM NaCl) used for pH titration experiments.

$^{13}C/^{15}N$ labeled H2A-unlabeled H2B	
Buffer type (20 mM)	pH
HAC	4.5
HAC	5.1
MES	5.6
MES	6.0
MES	6.5
Pi	6.8
Pi	7.5
Pi	7.9
CHES	9.1
unlabeled labeled H2A-^{15}N H2B	
HAC	4.4
HAC	5.0
MES	5.4
MES	5.8
MES	6.2
Pi	6.8
Pi	7.1
Pi	7.5

Table S5.3: all amide pKa and CSP values for H2A and H2B during the NMR pH titration. When the combined CSP is smaller than 10 Hz then no pKa value was calculated.

H2A backbone NH apparent pKa			
Residue number	Residue name	Combined CSP (Hz)	pKa
4	GLY	31	5.66
6	GLY	23	6.25
9	VAL	17	5.87
10	LYS	19	5.94
11	GLY	31	6.20
12	LYS	31	5.52
13	ALA	190	6.10
14	LYS	133	6.25
15	SER	120	6.13
17	SER	65	5.67
18	ASN	46	5.75
20	ALA	118	5.09
21	GLY	77	5.68
22	LEU	82	5.44
23	GLN	460	4.38
24	PHE	70	4.67
26	VAL	30	4.73
27	GLY	58	5.28
28	ARG	48	5.81
29	ILE	30	5.23
30	HIS	52	4.36
32	LEU	25	4.29
33	LEU	49	4.52
34	ARG	120	4.74
35	LYS	28	5.37
36	GLY	34	5.48
38	TYR	47	4.79
39	ALA	29	6.67
40	GLU	78	5.91
41	ARG	30	5.22
42	VAL	27	4.45

43	GLY	45	4.83
44	ALA	15	4.25
46	ALA	18	6.08
51	ALA	50	5.05
54	MET	62	5.72
55	GLU	60	5.63
56	TYR	23	6.50
57	LEU	32	4.47
58	ALA	65	5.86
59	ALA	118	6.14
60	GLU	63	6.15
61	VAL	110	5.80
62	LEU	80	5.54
63	GLU	120	5.52
64	LEU	55	4.96
65	ALA	110	5.75
66	GLY	145	5.64
67	ASN	150	5.72
69	ALA	85	5.55
70	ARG	79	5.60
71	ASP	13	4.58
72	ASN	18	5.58
73	LYS	24	6.41
76	ARG	33	6.02
77	ILE	23	7.11
80	ARG	14	5.53
81	HIS	13	6.14
82	LEU	27	5.30
83	GLN	17	5.62
84	LEU	22	3.53
85	ALA	25	5.69
86	ILE	42	4.71
87	ARG	41	4.59
88	ASN	30	4.85
89	ASP	87	4.96
90	GLU	15	4.24
91	GLU	45	5.36

92	LEU	140	5.82
93	ASN	110	4.91
94	LYS	140	5.01
95	LEU	50	4.74
99	VAL	25	4.11
100	THR	27	5.04
101	ILE	13	4.18
120	GLU	18	4.18
123	ALA	26	4.08
H2B backbone NH apparent pKa			
29	LYS	28	7.32
30	ARG	18	5.10
31	LYS	21	5.23
32	GLU	12	6.10
33	SER	23	5.18
35	ALA	53	4.91
36	ILE	25	5.31
37	TYR	22	4.48
40	LYS	32	6.29
41	VAL	46	5.81
42	LEU	21	5.94
43	LYS	48	6.00
44	GLN	40	5.59
45	VAL	33	6.46
46	HIS	140	5.77
48	ASP	50	5.83
49	THR	90	5.80
50	GLY	23	6.05
51	ILE	18	5.42
52	SER	17	4.67
55	ALA	90	4.96
56	MET	49	6.62
57	SER	15	4.82
58	ILE	16	5.85
59	MET	27	5.02
61	SER	40	5.52
62	PHE	45	4.38

63	VAL	40	3.02
67	PHE	20	3.75
69	ARG	46	3.42
70	ILE	28	4.75
71	ALA	13	6.44
74	ALA	33	5.40
76	ARG	39	6.79
77	LEU	16	5.88
79	HIS	100	6.91
80	TYR	22	6.52
81	ASN	35	6.42
82	LYS	52	5.81
83	ARG	16	5.23
87	THR	61	3.77
90	GLU	18	4.23
91	ILE	26	3.63
93	THR	40	6.28
94	ALA	30	5.06
95	VAL	29	5.03
97	LEU	38	5.21
98	LEU	23	4.54
102	GLU	300	4.50
103	LEU	70	4.19
104	ALA	140	5.74
105	LYS	95	5.70
106	HIS	71	6.61
108	VAL	125	5.84
109	SER	41	5.85
110	GLU	104	5.14
111	GLY	33	4.33
114	ALA	15	4.40
115	VAL	48	4.41
117	LYS	115	5.78
118	TYR	21	7.38
119	THR	114	5.81
120	SER	79	5.79
121	SER	45	5.74

122	LYS	21	4.25